

AMENDMENTS

In the claims:

Claim 1-14 (Cancelled).

15. (Currently Amended) A computer-based method for deriving a linear classifier for classifying a test gene expression dataset comprising:

- (a) providing a reference gene expression dataset comprising two subsets of data, wherein one subset is labeled in the class and the other subset is labeled outside the class; and
- (b) deriving a linear classification rule by reducing minimizing the value of a loss function associated withon said reference gene expression dataset, thereby deriving
and applying said a linear classification rule to a test gene expression dataset
thereby capable of determining the classification of the test gene expression dataset as in the class or outside the class.

16. (Currently Amended) The method of claim 15, wherein the reference gene expression dataset is a chemogenomic dataset based oncomprising gene expression levels measured in response to in vivo compound treatments.

17. (Original) The method of claim 15 wherein the type of loss function is selected from the group consisting of support vector machine, logistic regression, and minimax probability machine.

18. (Original) A computer ~~program product~~ readable medium comprising computer-executable code for deriving a linear classifier for classifying a test gene expression dataset, said code comprising instructions for:

- (a) computer code for querying accepting input of a reference gene expression dataset comprising two subsets of data, wherein one subset is labeled in the class and the other subset is labeled outside the class; and
- (b) computer code for deriving a linear classification rule by reducing minimizing the value of a loss function associated withon said reference gene expression dataset, thereby deriving a linear classifier capable of determining the classification of the test gene expression dataset as in the class or outside the class.;

~~computer code for applying said linear classification rule to a test gene expression dataset and thereby determining the classification of the test gene expression dataset; and~~

~~computer code for outputting the test dataset classification to the user.~~

19. (Original) The computer code product of claim 18 wherein the type of loss function is selected from the group consisting of support vector machine, logistic regression, and minimax probability machine.

20. (New) The method of claim 15, wherein minimizing the value of the loss function is performed according to the formulation:

$$\min_{w,b} \sum_{i=1}^n |w_i| + C \sum_{i=1}^N e_i$$

$$\text{subject to } y_i (w^T \cdot x_i + b) \geq 1 - e_i, \quad i = 1, \dots, N$$

$$\text{and wherein, } C = 1/\rho, \text{ and } \sigma = 1.$$

21. (New) The method of claim 15, wherein minimizing the value of the loss function comprises reducing a worse-case value of the loss function.

22. (New) The method of claim 15, wherein data labeled in the class represents a biological state or a biological activity.

23. (New) The method of claim 15, wherein data labeled in the class represents a biological state resulting from a compound treatment.

24. (New) The method of claim 15, wherein data labeled in the class represents a structural class of compounds.

25. (New) The computer code product of claim 18 wherein minimizing the value of the loss function is performed according to the formulation:

$$\min_{w,b} \sum_{i=1}^n |w_i| + C \sum_{i=1}^N e_i$$

subject to $y_i(w^T \cdot x_i + b) \geq 1 - e_i$, $i = 1, \dots, N$

and wherein, $C = 1/\rho$, and $\sigma = 1$.

26. (New) The computer code product of claim 18, wherein minimizing the value of the loss function comprises reducing a worse-case value of the loss function.
27. (New) The computer code product of claim 18, wherein the reference gene expression dataset is a chemogenomic dataset comprising gene expression levels measured in response to *in vivo* compound treatments.
28. (New) The computer code product of claim 18, wherein data labeled in the class represents a biological state or a biological activity.
29. (New) The computer code product of claim 18, wherein data labeled in the class represents a biological state resulting from a compound treatment.
30. (New) The computer code product of claim 18, wherein data labeled in the class represents a structural class of compounds.